

2067P-05

340001274

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出 願 年 月 日
Date of Application:

2000年 9月28日

出 願 番 号
Application Number:

特願2000-300563

出 願 人
Applicant(s):

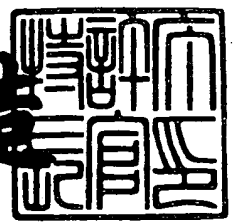
株式会社日立製作所
日立ソフトウェアエンジニアリング株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 4月 6日

特 許 庁 長 官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2001-3027787

【書類名】 特許願

【整理番号】 K00012741

【提出日】 平成12年 9月28日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06
G06F 3/301

【請求項の数】 9

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 井戸 健嗣

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 後藤 陽一

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 横畑 静生

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 本間 繁雄

【発明者】

【住所又は居所】 神奈川県横浜市中区尾上町 6 丁目 8 1 番地 日立ソフトウェアエンジニアリング株式会社内

【氏名】 吉野 利之

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【特許出願人】

【識別番号】 000233055

【氏名又は名称】 日立ソフトウェアエンジニアリング株式会社

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶制御装置

【特許請求の範囲】

【請求項 1】

複数コンピュータと該複数コンピュータとチャンネルパスで接続され、該チャンネルパスによって入出力を行なう記憶制御装置と、該記憶制御装置配下において、該コンピュータの入出力データを記憶する記憶デバイスとから構成されるコンピュータシステムに於いて、

該チャンネルパスが接続される該記憶制御装置の複数のチャンネルポートを優先チャンネルポートと非優先チャンネルポートに区別し、優先チャンネルポートに定義されたチャンネルポートは、該コンピュータからの I/O 処理を抑制することなく処理し、逆に非優先チャンネルポートに定義されたチャンネルポートに対しては、単位時間当たりの I/O 数などのような I/O 処理単位の目標値を定義し、該コンピュータからの I/O 処理を該 I/O 処理単位に近づける様にフィードバック制御しながら行なう手段を持ち、該手段によって該非優先チャンネルポートの I/O 処理が該優先チャンネルポートの I/O 処理に与える影響度を制御することを特徴とする記憶制御装置。

【請求項 2】

前記請求項 1 に記載の記憶制御装置に於いて、

該優先チャンネルポートに対して I/O 処理単位の目標値を定義し、該優先チャンネルポートの I/O 処理が、該 I/O 処理単位の目標値に近づく様に、該非優先チャンネルポートでの I/O 処理をフィードバック制御しながら行なう手段を持ち、該手段によって該非優先チャンネルポートの I/O 処理が該優先チャンネルポートの I/O 処理に与える影響度を制御することを特徴とする記憶制御装置。

【請求項 3】

前記請求項 1、又は請求項 2 に記載の記憶制御装置に於いて、

該優先チャンネルポートに設定されたチャンネルポートの I/O 頻度が閾値よりも低い時には、該非優先チャンネルポートの I/O 処理の抑制を行わない様にするための閾値を定義し、該非優先チャンネルポートの I/O 処理を抑制しない手段を持ち

、該記憶制御装置全体の I / O 処理性能を維持することを特徴とする記憶制御装置。

【請求項 4】

前記請求項 1、又は請求項 2、又は請求項 3 に記載の記憶制御装置に於いて、
該記憶制御装置に対して I / O 処理要求を行なうコンピュータ単位、または、
World Wide Name 等の様なコンピュータのバス単位に、優先ホストと非優先ホストを区別し、チャンネルポート間だけでなく同一チャンネルポート内でも、該非優先ホストの I / O 処理を、該目標 I / O 処理単位に近づける様にフィードバック制御しながら行なう手段を持ち、該手段によって該非優先ホストの I / O 処理が該優先ホストの I / O 処理に与える影響度を制御することを特徴とする記憶制御装置。

【請求項 5】

前記請求項 1、又は請求項 2、又は請求項 3 に記載の記憶制御装置に於いて、
該記憶制御装置内の I / O 処理を行なう記憶デバイス単位に優先デバイスと非優先デバイスを区別し、該非優先デバイスの I / O 処理を、該目標 I / O 処理単位に近づける様にフィードバック制御しながら行なう手段を持ち、該手段によって該非優先デバイスの I / O 処理が該優先デバイスの I / O 処理に与える影響度を制御することを特徴とする記憶制御装置。

【請求項 6】

前記請求項 5 に記載の記憶制御装置に於いて、
該記憶デバイス内の記憶エリアを優先エリアと非優先エリアに区別し、該非優先エリアの I / O 処理を、該目標 I / O 処理単位に近づける様にフィードバック制御しながら行なう手段を持ち、該手段によって該非優先エリアの I / O 処理が該優先エリアの I / O 処理に与える影響度を制御することを特徴とする記憶制御装置。

【請求項 7】

複数のコンピュータと接続する複数のポートと、前期コンピュータからの I / O を制御するコントローラとからなる記憶制御装置と、
前期記憶制御装置が受け取った前期コンピュータからの I / O を記憶する複数の

記憶デバイスで構成する記憶装置とからなる記憶システムであって、
前期コントローラは、予め前期ポート単位に優先または非優先のどちらか一方の値を保持した優先情報テーブルを具備し、前期優先情報テーブルで非優先の値を持ったポートが受信した I / O 処理の開始を、予め定義された時間遅らせることを特徴とした記憶システム。

【請求項 8】

複数のコンピュータと接続する複数のポートと、前期コンピュータからの I / O を制御するコントローラとからなる記憶制御装置と、
前期記憶制御装置が受け取った前期コンピュータからの I / O を記憶する複数の記憶デバイスで構成する記憶装置とからなる記憶システムであって、
前期コントローラは、予め前期コンピュータ単位に優先または非優先のどちらか一方の値を保持した優先情報テーブルを具備し、前期優先情報テーブルで非優先の値を持ったコンピュータから受信した I / O 処理の開始を、予め定義された時間遅らせることを特徴とした記憶システム。

【請求項 9】

複数のコンピュータと接続する複数のポートと、前期コンピュータからの I / O を制御するコントローラとからなる記憶制御装置と、
前期記憶制御装置が受け取った前期コンピュータからの I / O を記憶する複数の記憶デバイスで構成する記憶装置とからなる記憶システムであって、
前期コントローラは、予め前期記憶デバイス単位に優先または非優先のどちらか一方の値を保持した優先情報テーブルを具備し、前期優先情報テーブルで非優先の値を持った記憶デバイスに対する I / O 処理の開始を、予め定義された時間遅らせることを特徴とした記憶システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、記憶制御装置に関し、特に詳細には、コンピュータと該コンピュータと I / O 処理を行う記憶制御装置間の負荷バランス制御方式技術に関するものである。

【 0 0 0 2 】

【従来の技術】

近年の計算機システムでは、記憶制御装置に対する I / O 処理が全体性能のネックとなり易いために、コンピュータからの I / O に対して、記憶制御装置は、各 I / O 処理を該記憶制御装置で出し得る最高の性能で実行する事が必要になる。また、複数のコンピュータから複数のパスを通しての I / O 処理要求を同時に実行する必要もある。

【 0 0 0 3 】

一般的に、記憶制御装置の性能を向上する手段として、キャッシュメモリが使用され、LRU (Least Recently Used.) 制御などにより頻繁にアクセスされるデータを出来るだけキャッシュに残すことで、I / O 処理性能を向上させることができる。しかし、キャッシュ上に残っているデータが、必ずしも頻繁にアクセスされるデータになるとは限らず、I / O 処理におけるアクセスデータパターンによっては、キャッシュヒット率が上がらず、期待する I / O 処理性能が得られない場合がある。

【 0 0 0 4 】

この様な問題を解決する手段として、例えば、特開平 8 - 2 6 3 3 8 0 号公報で開示される技術では、ホストからのアクセス履歴を記録及び解析し、有効なデータをキャッシュ上に長く留め、キャッシュしても無意味と判断したデータは、キャッシュ上から積極的に排除することで、キャッシュヒット率を高め、I / O 処理性能を向上するとしている。

【 0 0 0 5 】

【発明が解決しようとする課題】

前記公報で開示される技術は、あくまでも使用可能なキャッシュメモリ資源が、十分に存在する場合に動作した時に有効な手段である。I / O 処理が高負荷になるにつれ、前記キャッシュメモリ資源は前記複数のパス間で使用し合うために、使用可能な該キャッシュメモリ資源を確保するための待ち時間が生じ、I / O 処理のスループットが低下する。また、複数のコンピュータが、同時に同一のキャッシュメモリ資源を必要とした場合にも該キャッシュメモリ資源を確保するた

めの待ち時間が生じ、業務等における I / O 処理優先度の高いコンピュータと低いコンピュータ間で該キャッシュメモリ資源確保の競合が発生していた場合、該 I / O 処理の優先度の低いコンピュータの処理が該 I / O 処理優先度の高いコンピュータの処理に影響を与えてしまう。

【 0 0 0 6 】

また、記憶制御装置全体としてのキャッシュメモリの使用率が高くなるにつれ、キャッシュメモリ資源を確保するための競合が発生し、例えば、バックグラウンドで行う様な処理優先度の比較的低い I / O が、例えば、オンラインで実行する様な処理優先度の高い I / O に影響を与えてしまうことになる。

【 0 0 0 7 】

本発明の目的は、複数のパスから複数のコンピュータが、複数の記憶デバイスに発行する I / O を処理する記憶制御装置に於いて、処理優先度の比較的低い I / O が、処理優先度の高い I / O に影響を与えないこと、及び記憶制御装置全体としての処理性能を維持することである。

【 0 0 0 8 】

【課題を解決するための手段】

上記目的を達成する為に、記憶制御装置内に設けたプロセッサ間共有メモリに於いて、チャネルポート単位毎に優先チャネルポート情報を持たせ、「優先」と設定されているチャネルポートに対しては目標 I / O P S (I / O p e r s e c o n d) 情報を定義し、「優先」ポートの I / O 処理の目標値が該目標 I / O P S に達することができる様に、「優先」と設定されていないチャネルポートに対しての I / O 処理に対してフィードバック抑制を実施する。

【 0 0 0 9 】

また、コンピュータ毎に優先ホスト情報を持たせ、「優先」と設定されているコンピュータに対して目標 I / O P S 情報を定義し、「優先」コンピュータからの I / O 処理の目標値が当該目標 I / O P S に達することができるよう、「優先」と設定されていないコンピュータからの I / O 処理に対してフィードバック抑制を実施する。

【 0 0 1 0 】

同様、デバイス毎に優先デバイス情報を持たせ、「優先」と設定されているデバイス対して目標 I / O P S 情報を定義し、「優先」デバイスへの I / O 処理の目標値が当該目標 I / O P S に達することができるよう、「優先」と設定されていないデバイスへの I / O 処理に対してフィードバック抑制を実施する。

【 0 0 1 1 】

【発明の実施の形態】

以下、図面を参照しながら本発明についての説明をする。

【 0 0 1 2 】

図 1 は、本発明におけるコンピュータシステムの一実施例を示した図である。図 1 のコンピュータシステムに於いて、記憶制御装置 3 0 1 は、チャンネルパス 2 0 1 ~ 2 0 4 を介してコンピュータ 1 0 1 ~ 1 0 4 に接続され、複数記憶デバイス 6 0 1、6 0 2、6 0 3、6 0 4 … に対して要求される I / O 処理を制御する。記憶制御装置 3 0 1 は、チャンネルポート 4 0 1 ~ 4 0 4 及び I / O 処理制御部 (プロセッサ) 5 0 1 ~ 5 0 4 と各々に備わる個別メモリ 5 0 6 ~ 5 0 9、及び I / O 処理制御部 (プロセッサ) 5 0 1 ~ 5 0 4 からアクセス可能な共有メモリ 5 0 5、キャッシュメモリ 5 1 0 からなる。I / O 処理制御部 5 0 1 ~ 5 0 4 は、それぞれ独立に複数記憶デバイス 6 0 1 … に対してデータの入出力を行う。共有メモリ 5 0 5 には、コンピュータ 1 0 1 ~ 1 0 4 からの I / O 処理要求を実行する際に I / O 処理の優先度を判定する為のチャンネルポート情報管理テーブル 5 1 1、ホスト情報管理テーブル 5 2 1、記憶デバイス情報管理テーブル 5 3 1 が格納される。毎回の I / O 処理開始時に、これらのテーブルを元に I / O 処理開始可否を決定する。チャンネルポート情報管理テーブル 5 1 1 では、処理を開始しようとしている I / O の対象チャンネルパスが「優先」かどうかをチェックし、優先であれば、そのまま処理を継続、優先でなければ、該チャンネルの I / O にある程度の抑制をかけながら、I / O 処理を実行する。同様に、ホスト情報管理テーブル 5 2 1 では、処理を開始しようとしている I / O の対象コンピュータが「優先」かどうかをチェックし、記憶デバイス情報管理テーブル 5 3 1 では、処理を開始しようとしている I / O の対象記憶デバイスが「優先」かどうかをチェックすること

で該 I / O の動作を制御する。

【 0 0 1 3 】

図 2 は、記憶制御装置 3 0 1 内のチャネルポート 4 0 1 に、コンピュータ 1 0 1 ~ 1 0 4 が全て接続されている構成を示したものである。この場合には、1 チャネルポートに I / O が集中するので、チャネルポート情報管理テーブル 5 1 1 の情報では、何の効果も及ぼさないが、コンピュータ 1 0 1 ~ 1 0 4 間で優先度を持たせたい場合には、ホスト情報管理テーブル 5 2 1 の設定によって、優先コンピュータを定義し、他コンピュータの I / O 処理の影響を抑えることができる。ホスト情報管理テーブル (5 2 1) の設定単位をコンピュータ単位ではなく World Wide Name 単位にした場合は、World Wide Name のパス単位に優先度をもたせることができる。

【 0 0 1 4 】

図 3 は、記憶制御装置 3 0 1 内のチャネルポート 4 0 1 に、コンピュータ 1 0 1 のみが接続されている構成を示したものである。この場合には、1 チャネルポートのみの I / O であり、また影響を及ぼし合うコンピュータもないため、チャネルポート情報管理テーブル 5 1 1 やホスト情報管理テーブル 5 2 1 の情報では、何の効果も及ぼさないが、記憶デバイス 6 0 1 ~ 6 0 4 … の間で優先度を持たせたい場合には、記憶デバイス情報管理テーブル 5 3 1 の設定によって、優先記憶デバイスを定義し、他記憶デバイスの I / O 処理の影響を抑えることができる。

【 0 0 1 5 】

図 4 は、共有メモリ 5 0 5 内にチャネルポート情報管理テーブル 5 1 1 を格納する図である。チャネルポート情報管理テーブル 5 1 1 は、チャネルポート毎に情報を管理し、各情報は、優先チャネルポート情報 5 1 2 と、非優先チャネルポート情報 5 1 6 に分かれる。優先チャネルポート情報 5 1 2 は、該チャネルポートが優先チャネルポートかどうかを判別する優先チャネルポート識別情報 P c 5 1 3 と、非優先チャネルポートが I / O 処理を抑制するかどうかを判断するための I O P S 閾値情報 I t 5 1 4 と、実際の当該チャネルポートの稼動状態を示す I O P S 稼動情報 I e 5 1 5 と、優先チャネルポートに設定する時に、目標とす

る I/O 処理性能を設定するための目標 IOPS 情報 I o 5 1 9 からなる。具体的には、IOPS 閾値情報 I t 5 1 4 は、現在稼働中の I/O 数より算出される IOPS 稼働情報 I e 5 1 5 と比較され、IOPS 稼働情報 I e 5 1 5 が IOPS 閾値情報 I t 5 1 4 を超えた場合、「優先」と設定されていないチャネルポートに対する I/O 処理が抑制されるようになる。

【 0 0 1 6 】

また、非優先チャネルポート情報 5 1 6 は、該チャネルポートが非優先チャネルポートの場合に I/O 処理を抑制するためのディレイ設定値 d t 5 1 7、ディレイタイマ D t 5 1 8 と、I/O 処理の上限を設定するための IOPS 上限値 I u 5 1 A、I/O 処理を抑制する時の下限を設定する IOPS 下限値 I_L 5 1 B、IOPS 稼働情報 I e 5 1 C からなる。具体的には、上述の IOPS 稼働情報 I e 5 1 4 が IOPS 閾値情報 I t 5 1 4 を超えた場合に、非優先チャネルポートから受け付けた I/O の処理開始をディレイ設定値 d t 5 1 7 分遅らせる（I/O 処理を待機させる）のに用いる。タイマ D t 5 1 8 は、ディレイ設定値 d t 5 1 7 の時間をカウントアップする為の領域で、非優先チャネルポートから受け付けた時点で値が初期化され、タイマのカウントアップを開始する。

【 0 0 1 7 】

図 5 は、共有メモリ 5 0 5 内にホスト情報管理テーブル 5 2 1 を格納する図である。ホスト情報管理テーブル 5 2 1 はホスト毎に情報を管理し、各情報は、優先ホスト情報 5 2 2 と、非優先ホスト情報 5 2 6 に分かれる。優先ホスト情報 5 2 2 は、該ホストが優先ホストかどうかを判別する優先ホスト識別情報 P h 5 2 3 と、非優先ホストが I/O 処理を抑制するかどうかを判断するための IOPS 閾値情報 I t 5 2 4 と、実際の当該ホストの稼働状態を示す IOPS 稼働情報 I e 5 2 5 と、優先ホストに設定する時に、目標とする I/O 処理性能を設定するための目標 IOPS 情報 I o 5 2 9 からなる。また、非優先ホスト情報 5 2 6 は、該ホストが非優先ホストの場合に I/O 処理を抑制するためのディレイ設定値 d t 5 2 7、ディレイタイマ D t 5 2 8 と、I/O 処理の上限を設定するための IOPS 上限値 I u 5 2 A、I/O 処理を抑制する時の下限を設定する IOPS 下限値 I_L 5 2 B、IOPS 稼働情報 I e 5 2 C からなる。

【 0 0 1 8 】

図 6 は、共有メモリ 5 0 5 内に記憶デバイス情報管理テーブル 5 3 1 を格納する図である。記憶デバイス情報管理テーブル 5 3 1 は記憶デバイス毎に情報を管理し、各情報は、優先記憶デバイス情報 5 3 2 と、非優先記憶デバイス情報 5 3 6 に分かれる。優先記憶デバイス情報 5 3 2 は、該記憶デバイスが優先記憶デバイスかどうかを判別する優先記憶デバイス識別情報 $Pd\ 5\ 3\ 3$ と、非優先記憶デバイスが I/O 処理を抑制するかどうかを判断するための IOPS 閾値情報 $It\ 5\ 3\ 4$ と、実際の当該記憶デバイスの稼動状態を示す IOPS 稼動情報 $Ie\ 5\ 3\ 5$ と、優先記憶デバイスに設定する時に、目標とする I/O 処理性能を設定するための目標 IOPS 情報 $Io\ 5\ 3\ 9$ からなる。また、非優先記憶デバイス情報 5 3 6 は、該記憶デバイスが非優先記憶デバイスの場合に I/O 処理を抑制するためのディレイ設定値 $dt\ 5\ 3\ 7$ 、ディレイタイマ $Dt\ 5\ 3\ 8$ と、I/O 処理の上限を設定するための IOPS 上限値 $Iu\ 5\ 3\ A$ 、I/O 処理を抑制する時の下限を設定する IOPS 下限値 $IL\ 5\ 3\ B$ 、IOPS 稼動情報 $Ie\ 5\ 3\ C$ からなる。

【 0 0 1 9 】

図 7 は、チャネルポートをパラメータとして、I/O 処理におけるコマンド起動可否をチェックするフローを示した図である。図 7 のフローにおいて、まず優先チャネルポート識別情報 Pc を取得する 8 0 1。自チャネルポートが優先チャネルポートならばコマンド処理 OK で処理を終了する 8 0 2。自チャネルポートが優先チャネルポートでないならば、他ポートに優先チャネルポートが存在するかどうかをチェックし、優先チャネルポートが存在しなければ、コマンド処理 OK で処理を終了する 8 0 3。他ポートに優先チャネルポートが存在する場合は、優先チャネルポートの IOPS 閾値情報 It と IOPS 稼動情報 Ie を取得し 8 0 4、8 0 5、他チャネルポートに $It < Ie$ のポートがなければ、コマンド処理 OK で処理を終了する 8 0 6。 $It < Ie$ のポートがある場合、自チャネルポートのディレイ設定値 dt とディレイタイマ Dt を取得し 8 0 7、8 0 8、 $dt < Dt$ となる迄 Dt を更新しながら 8 1 1、コマンド処理を NG とする 8 1 3。 $dt < Dt$ となった時点で 8 0 9 Dt は初期化し 8 1 0、コマンド処理 OK とす

る 8 1 2。この様に非優先チャネルポートは自 I / O を抑制することによって、優先チャネルポートの I / O に影響を与えない様にする。

【 0 0 2 0 】

図 8 は、ホストをパラメータとして、I / O 処理におけるコマンド起動可否をチェックするフローを示した図である。図 8 のフローに於いて、まず優先ホスト識別情報 P h を取得する 9 0 1。自ホストが優先ホストならばコマンド処理 O K で処理を終了する 9 0 2。自ホストが優先ホストでないならば、他ホストに優先ホストが存在するかどうかをチェックし、優先ホストが存在しなければ、コマンド処理 O K で処理を終了する 9 0 3。他ホストに優先ホストが存在する場合は、優先ホストの I O P S 閾値情報 I t と I O P S 稼動情報 I e を取得し 9 0 4、9 0 5、他ホストに $I t < I e$ のホストがなければ、コマンド処理 O K で処理を終了する 9 0 6。 $I t < I e$ のホストがある場合、自ホストのディレイ設定値 d t とディレイタイマ D t を取得し 9 0 7、9 0 8、 $d t < D t$ となる迄 D t を更新しながら 9 1 1、コマンド処理を N G とする 9 1 3。 $d t < D t$ となった時点で 9 0 9 D t は初期化し 9 1 0、コマンド処理 O K とする 9 1 2。このように非優先ホストは自 I / O を抑制することによって、優先ホストの I / O に影響を与えないようにする。

【 0 0 2 1 】

同様に図 9 は、記憶デバイスをパラメータとして、I / O 処理におけるコマンド起動可否をチェックするフローを示した図である。図 9 のフローにおいて、まず優先記憶デバイス識別情報 P d を取得する 1 0 0 1。自記憶デバイスが優先記憶デバイスならばコマンド処理 O K で処理を終了する 1 0 0 2。自記憶デバイスが優先記憶デバイスでないならば、他記憶デバイスに優先記憶デバイスが存在するかどうかをチェックし、優先記憶デバイスが存在しなければ、コマンド処理 O K で処理を終了する 1 0 0 3。他記憶デバイスに優先記憶デバイスが存在する場合は、優先記憶デバイスの I O P S 閾値情報 I t と I O P S 稼動情報 I e を取得し 1 0 0 4、1 0 0 5、他記憶デバイスに $I t < I e$ の記憶デバイスがなければ、コマンド処理 O K で処理を終了する 1 0 0 6。 $I t < I e$ の記憶デバイスがある場合、自記憶デバイスのディレイ設定値 d t とディレイタイマ D t を取得し

1 0 0 7、1 0 0 8、 $dt < Dt$ となる迄 Dt を更新しながら 1 0 1 1、コマンド処理を NG とする 1 0 1 3。 $dt < Dt$ となった時点で 1 0 0 9 Dt は初期化し 1 0 1 0、コマンド処理 OK とする 1 0 1 2。このように非優先記憶デバイスは自 I/O を抑制することによって、優先記憶デバイスの I/O に影響を与えないようにする。

【 0 0 2 2 】

更に図 4 ～ 9 と同様にして、記憶デバイス内のエリアを優先エリアと非優先エリアに分割することで、同一記憶デバイス内であっても、非優先エリアに対する I/O 処理を抑制し、「優先」に指定されたエリアに対する I/O 処理に影響を与えないように制御する方法もある。

【 0 0 2 3 】

図 1 0 は、チャネルポートによるコマンド起動チェック (8 1 4) とホストによるコマンド起動チェック (9 1 4) と記憶デバイスによるコマンド起動チェック (1 0 1 4) を組み合わせた例である。図 1 0 の例では、チャネルポート、ホスト、記憶デバイス全てに「優先」の設定がされている場合、I/O 処理における全てのパラメータが「優先」を満たしている場合のみ、I/O 処理に抑制がかからない。それ以外の I/O 処理は、前記の I/O 処理に影響を及ぼさないようにするために、必ずディレイが生じることになる。また、コマンド処理 (1 1 0 4) ではコマンドに従ったリードまたはライト処理以外に、次の I/O 処理を想定して目的レコード以外の先読み処理を行なう場合などがあるが、このような処理に対してもチャネルポート、ホスト、記憶デバイス全てに「優先」の設定がされていなければ、当該先読み処理に抑制をかける。

【 0 0 2 4 】

図 1 1 は、記憶制御装置と接続されるサービスプロセッサ 7 0 1 から情報を設定し、非優先チャネルポートの I/O 処理のディレイを調整することで、優先チャネルポートの IOPS が目標 IOPS に最も近づく様に制御する為のフローを示した図である。まず、サービスプロセッサ 7 0 1 から優先チャネルポート識別情報 5 1 3、IOPS 閾値情報 5 1 4、目標 IOPS 情報 5 1 9 と、非優先チャネルポートの IOPS 上限値 5 1 A、IOPS 下限値 5 1 B を設定する。IOP

S 上限値 5 1 A の値は、優先チャネルポートの I / O 処理にどの程度の影響を及ぼすか正確にはわからないので、実環境に即した推測値を設定する。I / O 処理制御部では、設定された I O P S 上限値を取得し 1 1 1 1、I O P S 上限値から I / O 処理ディレイ設定値 d t 5 1 7 を算出し 1 1 1 2 この値を元に一定時間 I / O 処理を行う 1 1 1 3。その後優先チャネルポートの I O P S 稼動情報 I e 5 1 5、目標 I O P S 情報 I o 5 1 9、非優先チャネルポートの I O P S 下限値 I L 5 1 B を取得する 1 1 1 4、1 1 1 5。I e と I o に差分がなければ（または無視できる程度に差分が小さいと判断した場合）、優先チャネルポートの I / O 処理性能は目標に到達していると判断し、ディレイ設定値 5 1 7 はそのまま I / O 処理を継続する。優先チャネルポートの I / O 処理性能が目標に到達できていなくても自ポートの I / O P S が I u と I L の範囲から外れた場合にも調整は不可能と判断し、ディレイ設定値 d t 5 1 7 はそのまま I / O 処理を継続する。それ以外の場合には図に示す式に基づいてディレイ設定値 d t を再設定し、処理 1 1 1 3 へ戻る。これを繰り返すことにより、最適なディレイ設定値 d t を求めることができる。

【 0 0 2 5 】

優先 / 非優先ホスト、優先 / 非優先記憶デバイスの情報についても優先 / 非優先チャネルポートの情報と同様に設定することができる。

【 0 0 2 6 】

図 1 2 は、記憶制御装置と接続されるサービスプロセッサ 7 0 1 から設定した情報を基に、非優先チャネルポートの I / O 処理のディレイを調整し、非優先チャネルポートの I / O P S を非優先チャネルポートの I O P S 上限値に近づく様に制御して、非優先チャネルポートの I / O 処理を制限することで、優先チャネルポートの I / O P S が目標 I / O P S に最も近づくようにする為のフローを示した図である。まず、サービスプロセッサ 7 0 1 から優先チャネルポート識別情報 5 1 3、I O P S 閾値情報 5 1 4 と、非優先チャネルポートの I O P S 上限値 I u 5 1 A を設定する。I O P S 上限値 I u 5 1 A の値は、優先チャネルポートの I / O 処理にどの程度の影響を及ぼすか正確にはわからないので、実環境に即した推測値を設定する。I / O 処理制御部では、設定された I O P S 上限値を取

得し 1 1 2 1、IOPS 上限値から I/O 処理のディレイ設定値 $d t 5 1 7$ を算出し 1 1 2 2 この値を元に一定時間 I/O 処理を行う 1 1 2 3。その後優先チャネルポートの IOPS 稼動情報 $I e 5 1 C$ を取得する 1 1 2 4。 $I e$ と $I u$ に差分がなければ（または無視できる程度に差分が小さいと判断した場合）、優先チャネルポートの I/O 処理性能は目標に到達していると判断し、ディレイ設定値 $d t 5 1 7$ はそのまま I/O 処理を継続する。

【 0 0 2 7 】

それ以外の場合には図に示す式に基づいてディレイ設定値 $d t$ を再設定し 1 1 2 6、処理 1 1 2 3 へ戻る。これを繰り返すことにより、最適なディレイ設定値 $d t$ を求めることができる。この結果、優先チャネルポートの I/O P S が目標とする値に達していなければ、サービスプロセッサ 7 0 1 から I/O P S 上限値 $I u 5 1 A$ の値を再設定することを繰り返すことにより、最適な I/O 処理状態を模索する。優先／非優先ホスト、優先／非優先記憶デバイスの情報についても優先／非優先チャネルポートの情報と同様に設定することができる。

【 0 0 2 8 】

以上の方式により、優先的に行ないたい I/O の処理性能を最大限に引き出すためには、その他の優先度を落としてもよい I/O に制限をかけながら処理させることで実現することができる。

【 0 0 2 9 】

【発明の効果】

本発明による、優先度を落としてもよい I/O は抑制しながら処理させることで、優先的に行ないたい I/O の処理性能を維持する方式においては、以下の効果がある。

【 0 0 3 0 】

(1) 複数チャネルポートで I/O 処理を行なっている場合、性能を維持したいチャネルポートを優先チャネルポートとして設定しておくことで、記憶制御装置全体で I/O 処理が過負荷状態になっても、優先チャネルポートとして設定したチャネルポートは、一定の性能を維持することが可能である。

【 0 0 3 1 】

(2) 複数ホストが一つのチャネルポートに集中している場合には、優先したいホストを優先ホストとして設定しておくことで、記憶制御装置全体で I / O 処理が過負荷状態になっても、優先ホストとして設定したホストは、一定の性能を維持することが可能である。

【 0 0 3 2 】

(3) 一つのホストが複数の記憶デバイスに I / O 処理を行っている場合に、優先したい記憶デバイスを優先記憶デバイスとして設定しておくことで、記憶制御装置全体で I / O 処理が過負荷状態になっても、優先記憶デバイスとして設定した記憶デバイスへの I / O 処理は、一定の性能を維持することが可能である。

【 0 0 3 3 】

(4) 上記 (1) ~ (3) を組み合わせることで、特定のチャネルパスの特定のホストから特定の記憶デバイスへ発行される I / O のみ性能を維持するなど、優先する I / O 処理条件を特定化させることが可能になり、より細かい I / O 処理環境を構築することが可能である。

【図面の簡単な説明】

【図 1】

サービスプロセッサ及び、これと通信を行なうことができる複数のプロセッサを持つ記憶制御装置配下の複数の記憶デバイスと I / O 処理を行なうコンピュータシステムの概略を示す図

【図 2】

図 1 において、複数のコンピュータが一つのチャネルポートに接続され、記憶制御装置配下の複数の記憶デバイスと I / O 処理を行なうコンピュータシステムの概略を示す図

【図 3】

図 1 において、一つのコンピュータが記憶制御装置配下の複数の記憶デバイスと I / O 処理を行なうコンピュータシステムの概略を示す図

【図 4】

記憶制御装置内の共有メモリにあって、チャネルポートの優先 / 非優先を管理

、及び制御するためのテーブル

【図 5】

記憶制御装置内の共有メモリにあって、ホストの優先／非優先を管理、及び制御するためのテーブル

【図 6】

記憶制御装置内の共有メモリにあって、記憶デバイスの優先／非優先を管理、及び制御するためのテーブル。

【図 7】

チャンネルポートに関しての優先／非優先を判断し、コマンド起動可否をチェックする処理を示したフロー。

【図 8】

ホストに関しての優先／非優先を判断し、コマンド起動可否をチェックする処理を示したフロー。

【図 9】

記憶デバイスに関しての優先／非優先を判断し、コマンド起動可否をチェックする処理を示したフロー

【図 1 0】

チャンネルポート、ホスト、記憶デバイスに関しての優先／非優先をそれぞれ判断し、コマンド起動可否をチェックする処理を示したフロー

【図 1 1】

サービスプロセッサから優先／非優先のチャンネルポート情報を設定し、該情報を元に、「優先」に設定されたチャンネルポートの I / O 処理に与える影響を最小限に抑えるように、「優先」に設定されたチャンネルポートの目標値を基準に、「非優先」に設定されたチャンネルポートの I / O に対するディレイを調整する処理を示したフロー

【図 1 2】

サービスプロセッサから優先／非優先のチャンネルポート情報を設定し、該情報を元に、「優先」に設定されたチャンネルポートの I / O 処理に与える影響を最小限に抑えるように、「非優先」に設定されたチャンネルポートの上限値を基準に、

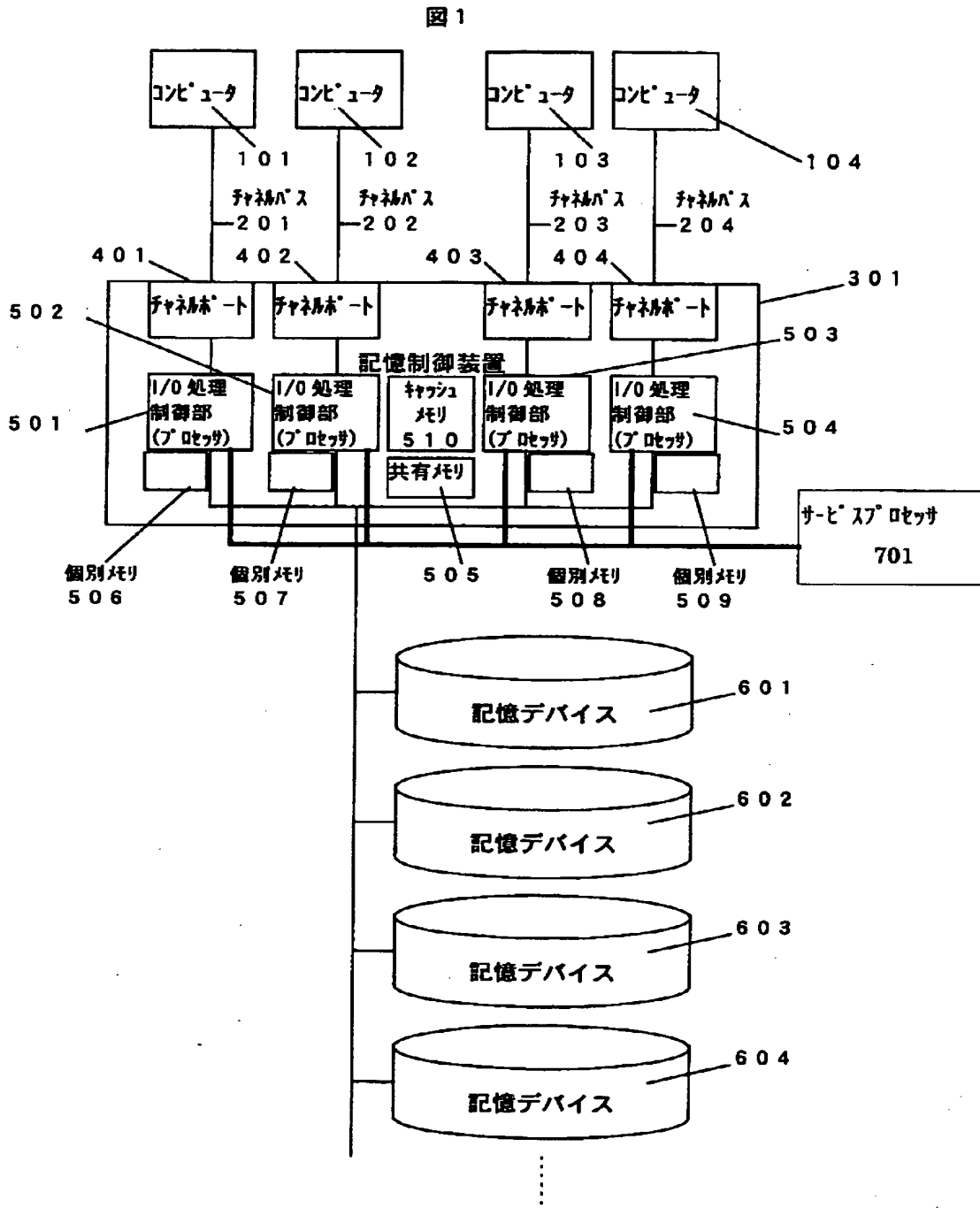
「非優先」に設定されたチャネルポートのI/Oに対するディレイを調整する処理を示したフロー

【符号の説明】

101～104…コンピュータ、201～204…チャネルパス、301～記憶制御装置、401～404…記憶制御装置のチャネルポート、501～504…I/O処理制御部（プロセッサ）、505…プロセッサ間共有メモリ、506～509…プロセッサ個別メモリ、510…キャッシュメモリ、601～604…記憶装置、701…サービスプロセッサ、511…チャネルポート情報管理テーブル、512…優先チャネルポート情報、513…優先チャネルポート識別情報、514…IOPS閾値情報、515…IOPS稼動情報、516…非優先チャネルポート情報、517…ディレイ設定値、518…ディレイタイマ、519…目標IOPS情報、51A…OPS上限値、51B…IOPS下限値、51C…IOPS稼動情報、521…ホスト情報管理テーブル、522…優先ホスト情報、523…優先ホスト識別情報、524…IOPS閾値情報、525…IOPS稼動情報、526…非優先ホスト情報、527…ディレイ設定値、528…ディレイタイマ、529…目標IOPS情報、52A…IOPS上限値、52B…I/OPS下限値、52C…IOPS稼動情報、531…記憶デバイス情報管理テーブル、532…優先記憶デバイス情報、533…優先記憶デバイス識別情報、534…IOPS閾値情報、535…IOPS稼動情報、536…非優先記憶デバイス情報、537…ディレイ設定値、538…ディレイタイマ、539…目標IOPS情報、53A…IOPS上限値、53B…IOPS下限値、53C…IOPS稼動情報、701…サービスプロセッサ、801～814、901～914、1001～1014、1101～1105、1111～1117、1121～1126…処理ステップ

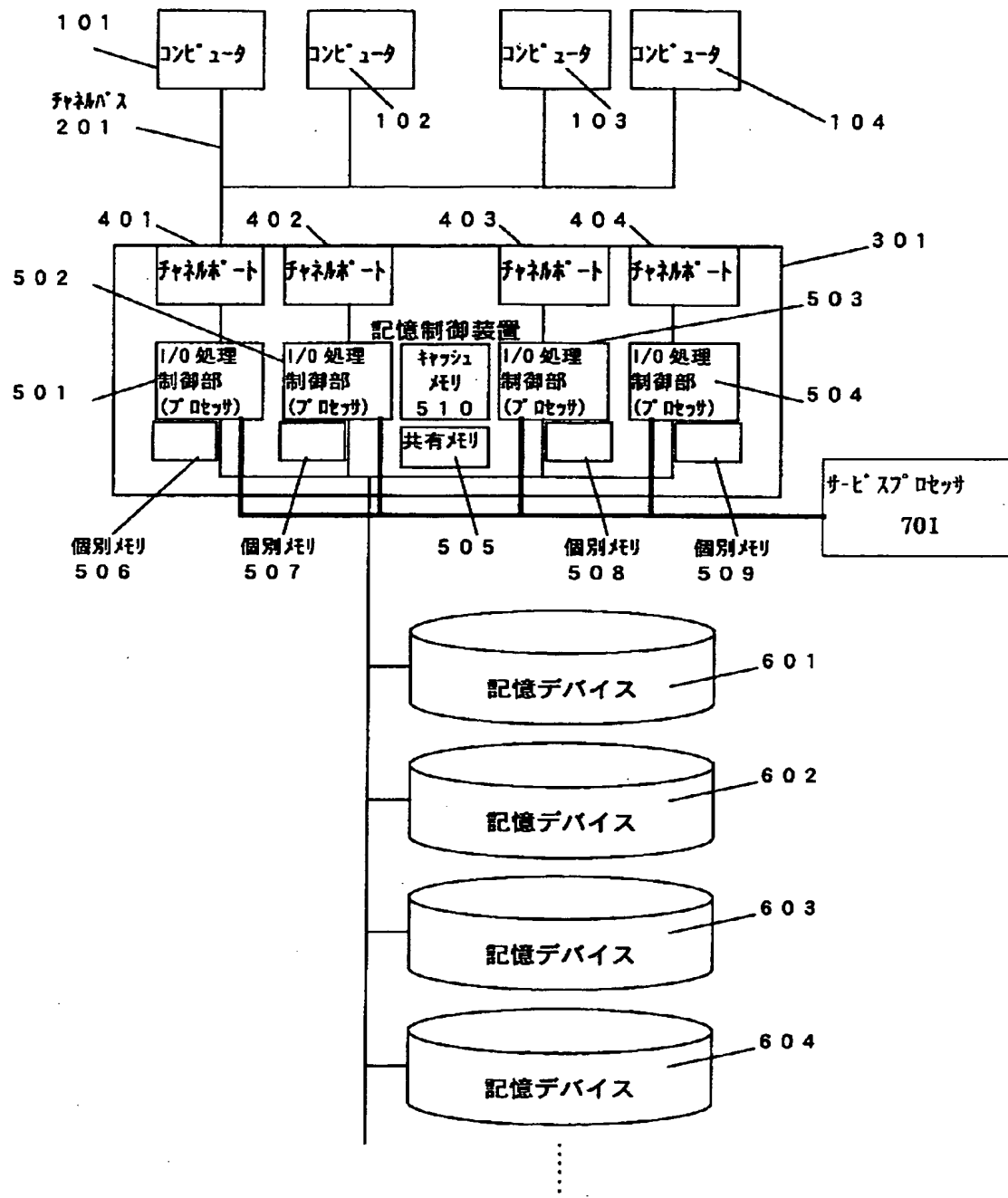
【書類名】 図面

【図1】



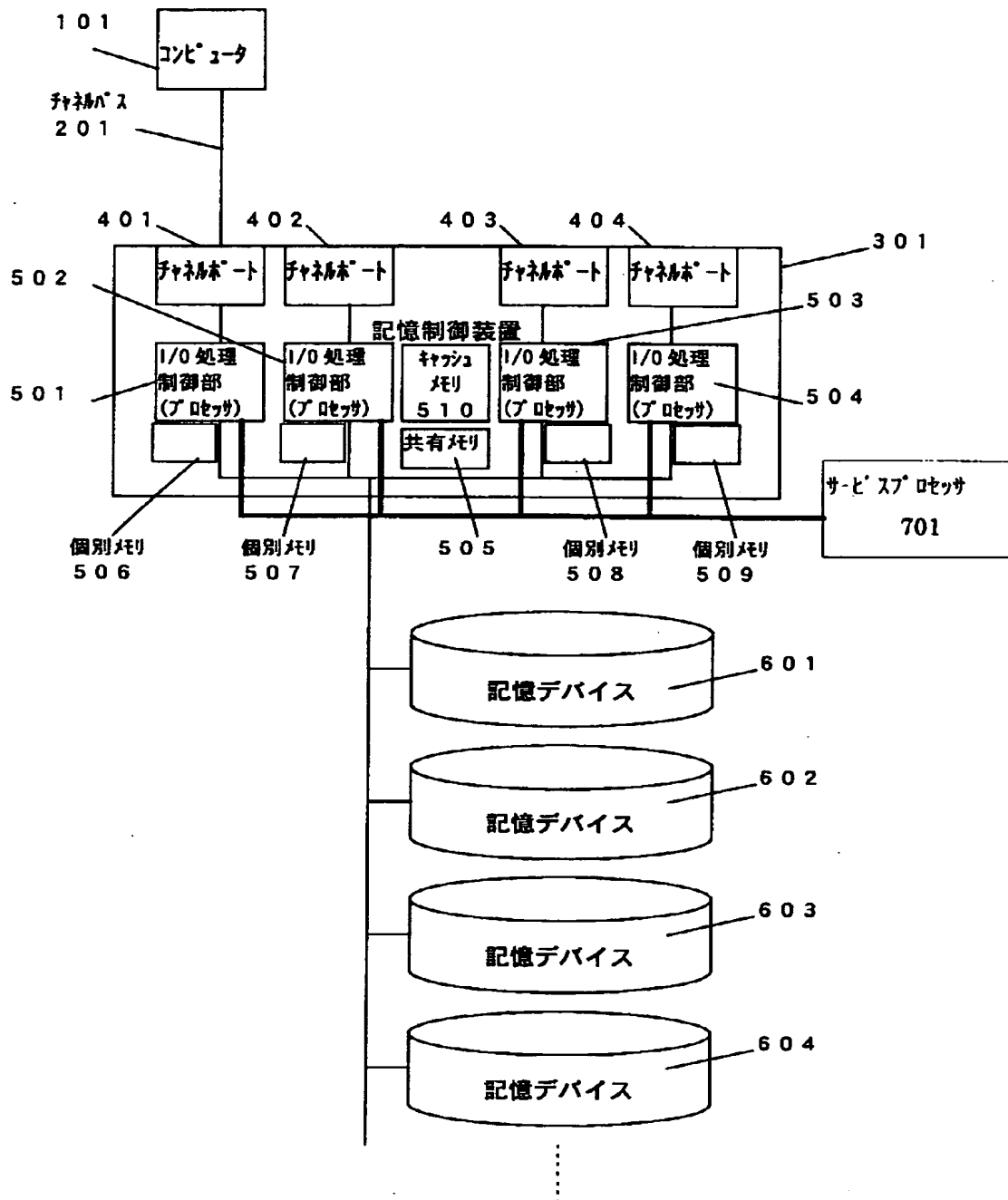
【図 2】

図 2

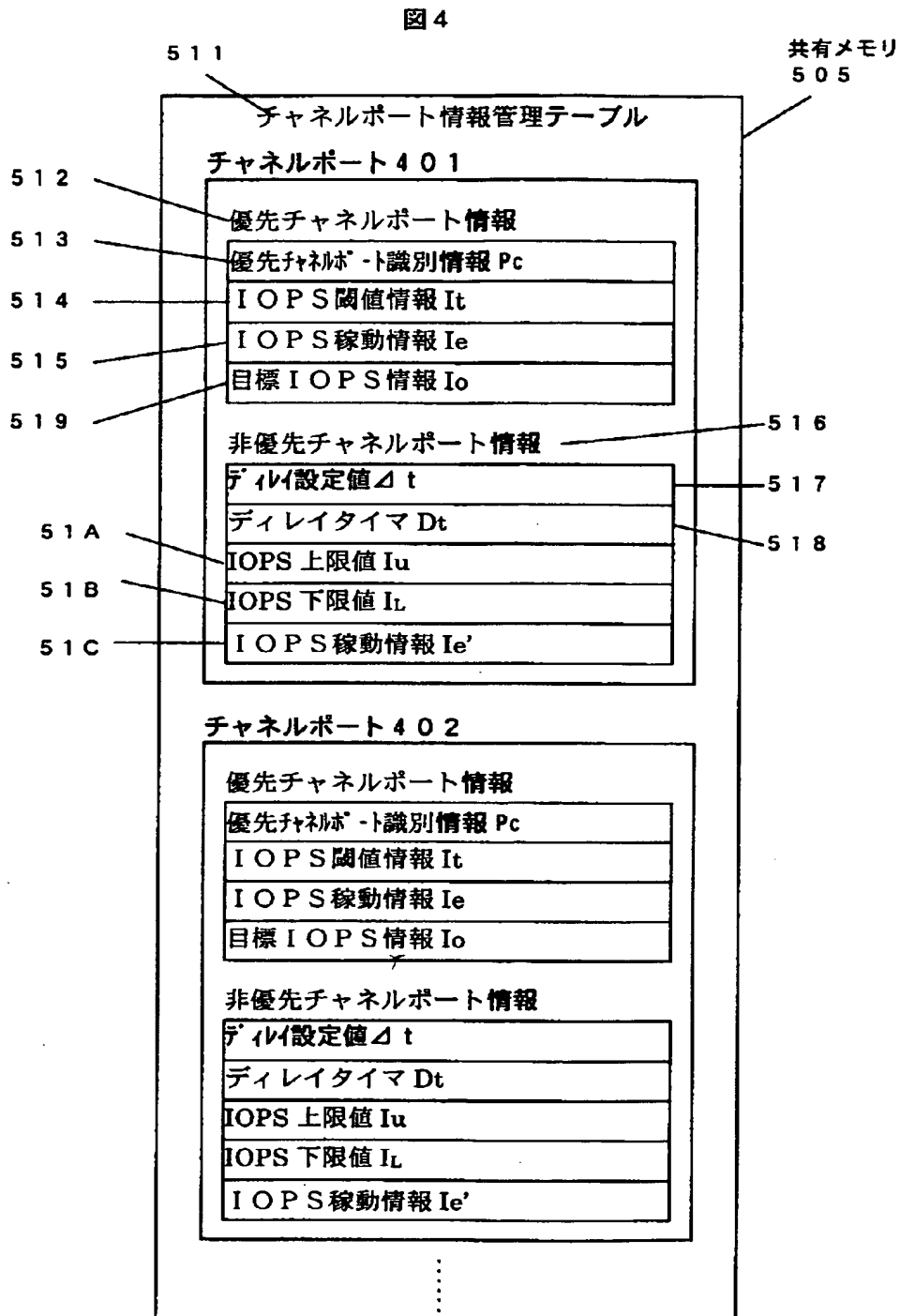


【図 3】

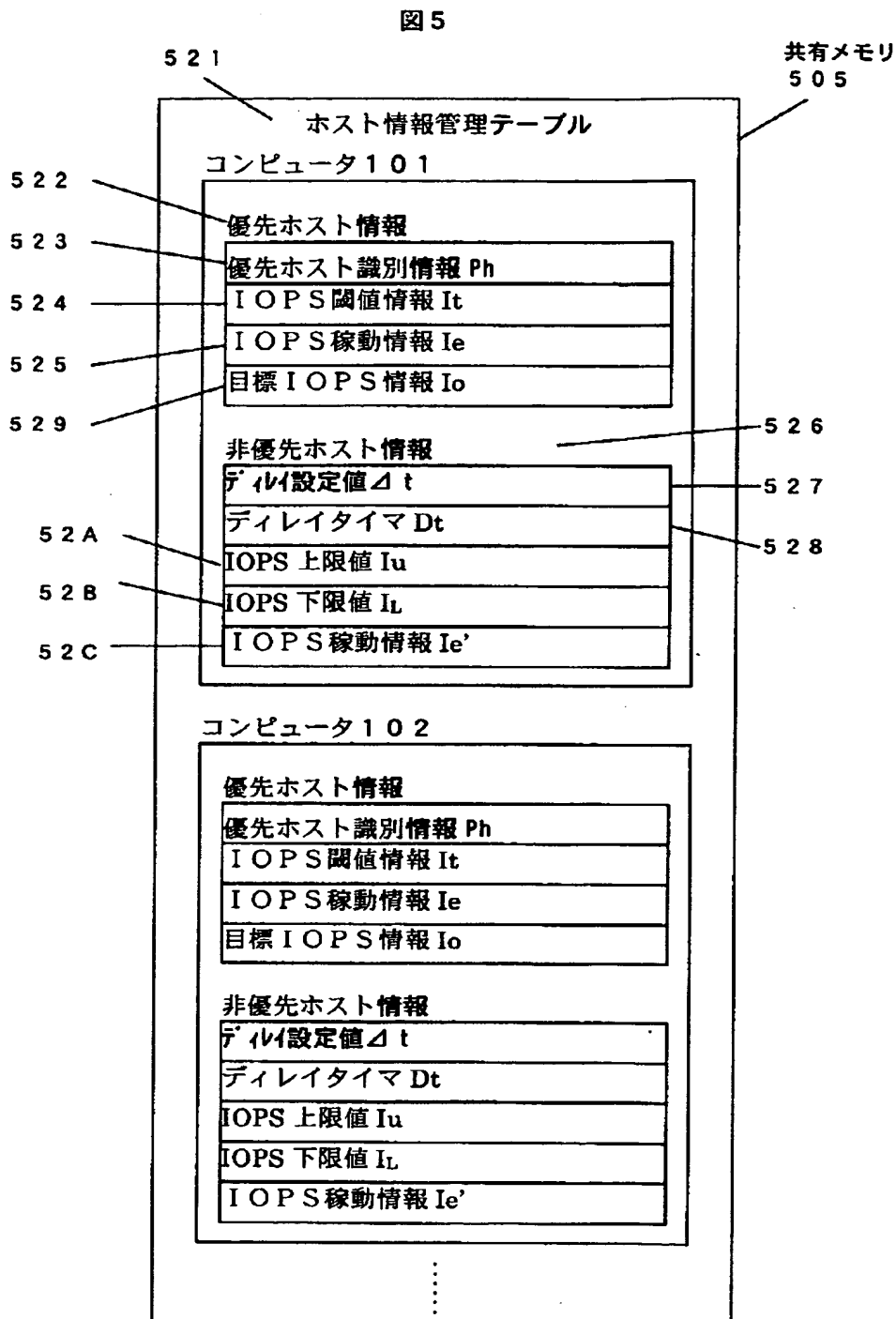
図 3



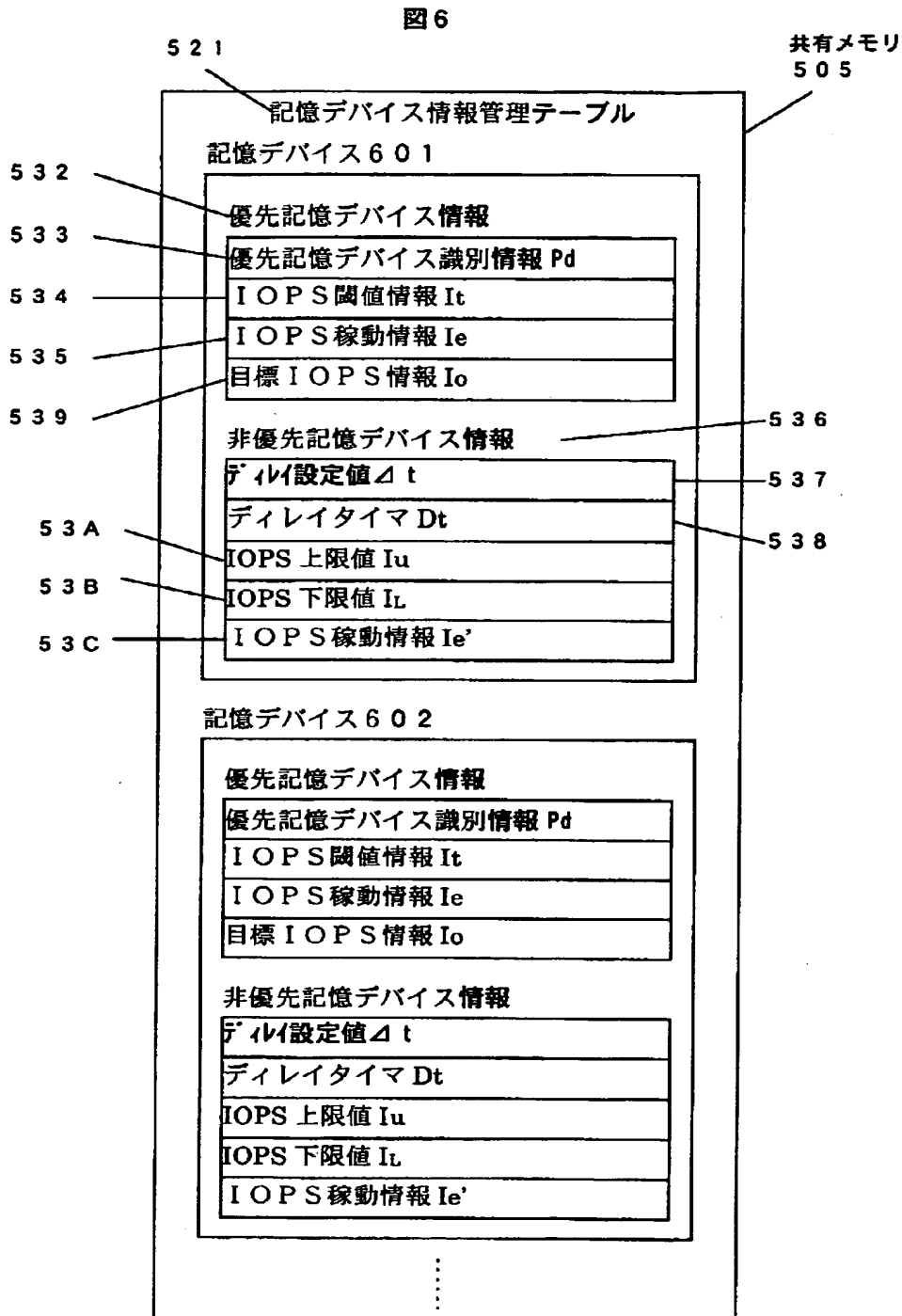
【図 4】



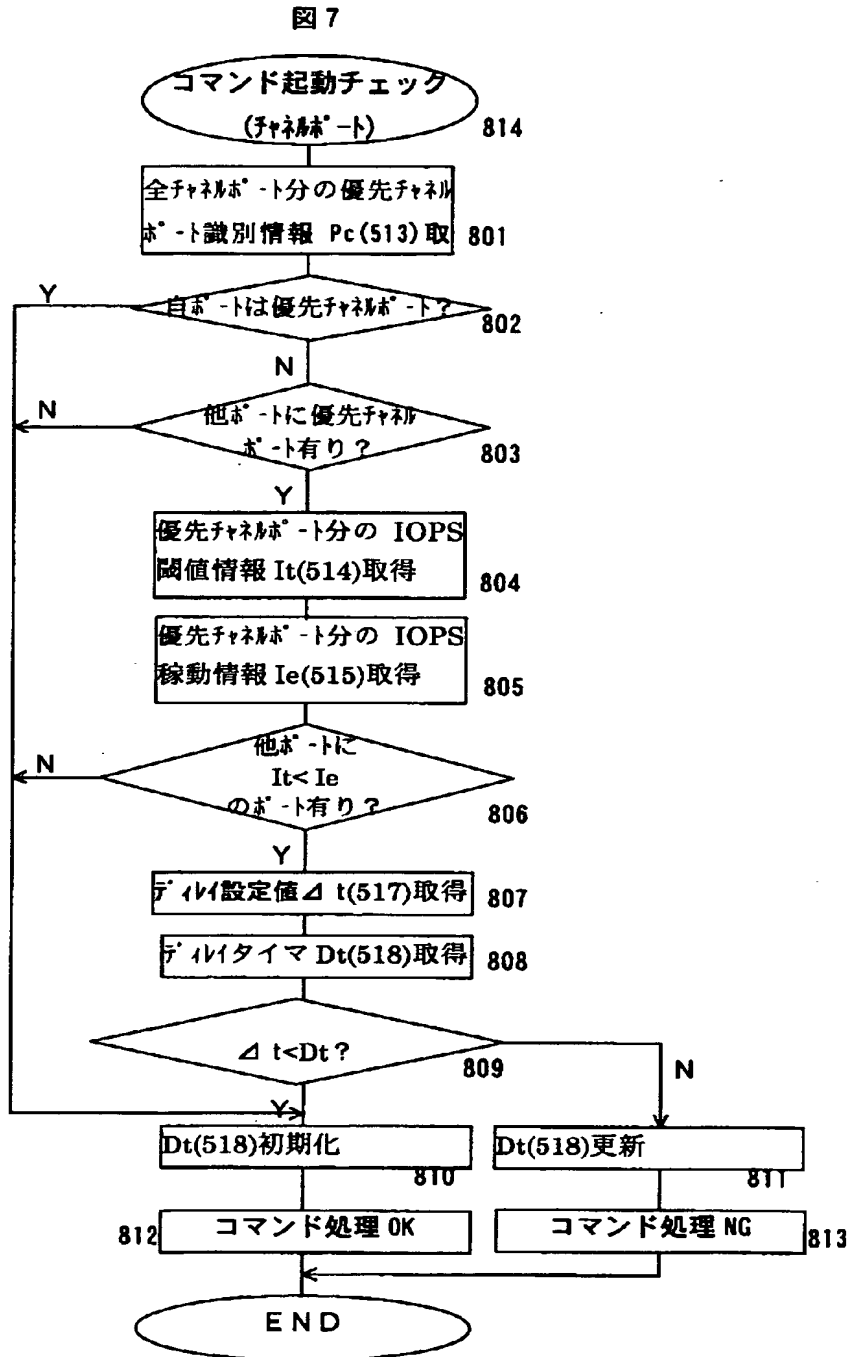
【図 5】



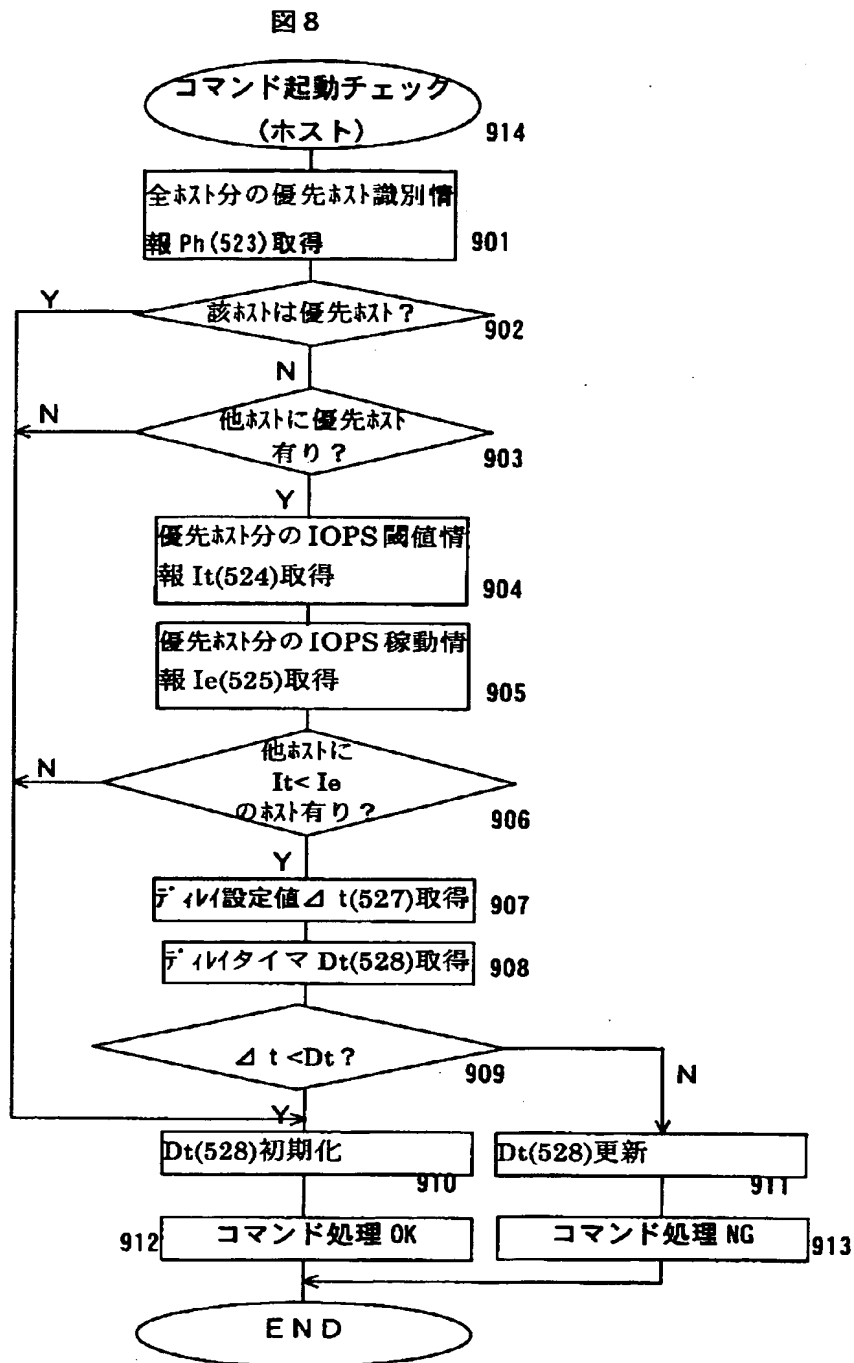
【図 6】



【図 7】

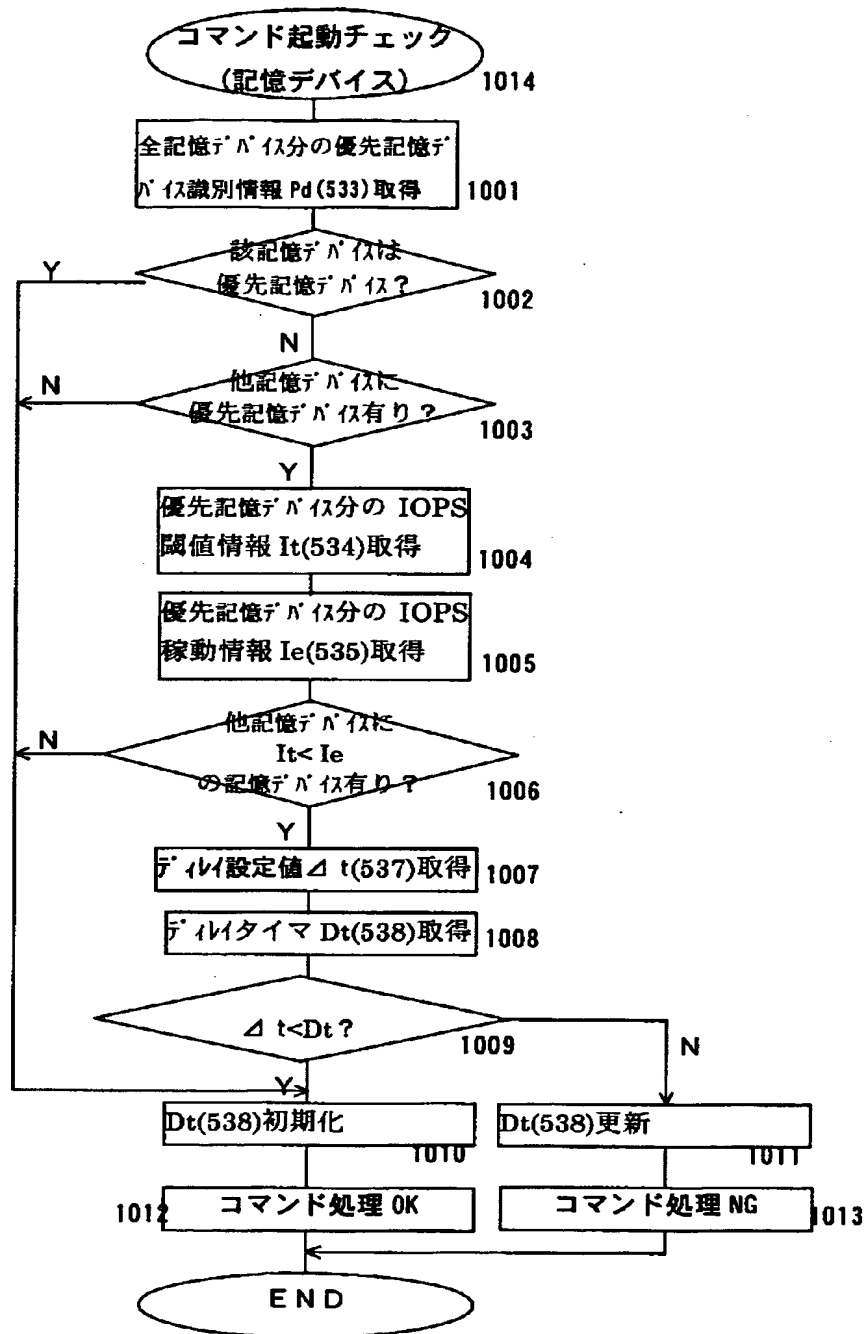


【図 8】



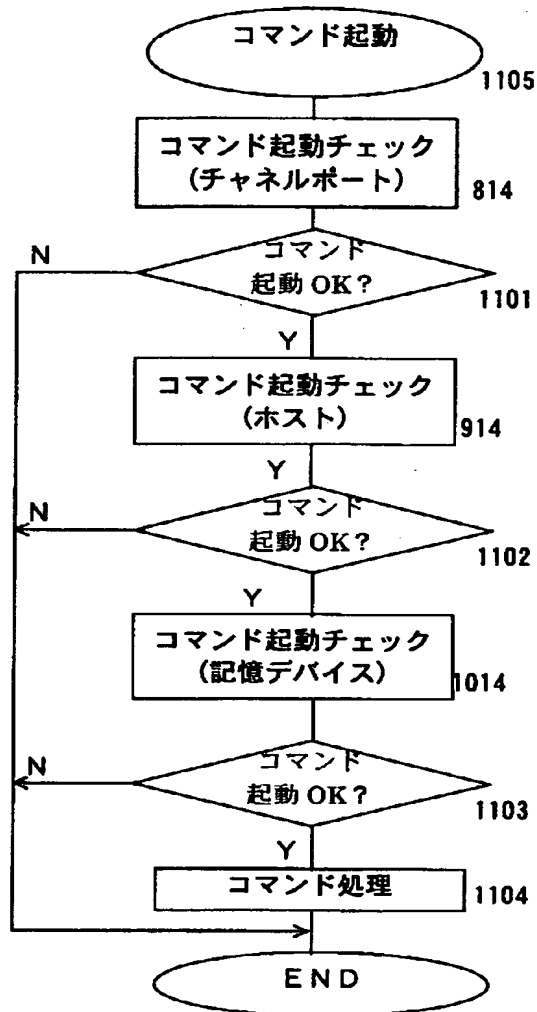
【図9】

図9



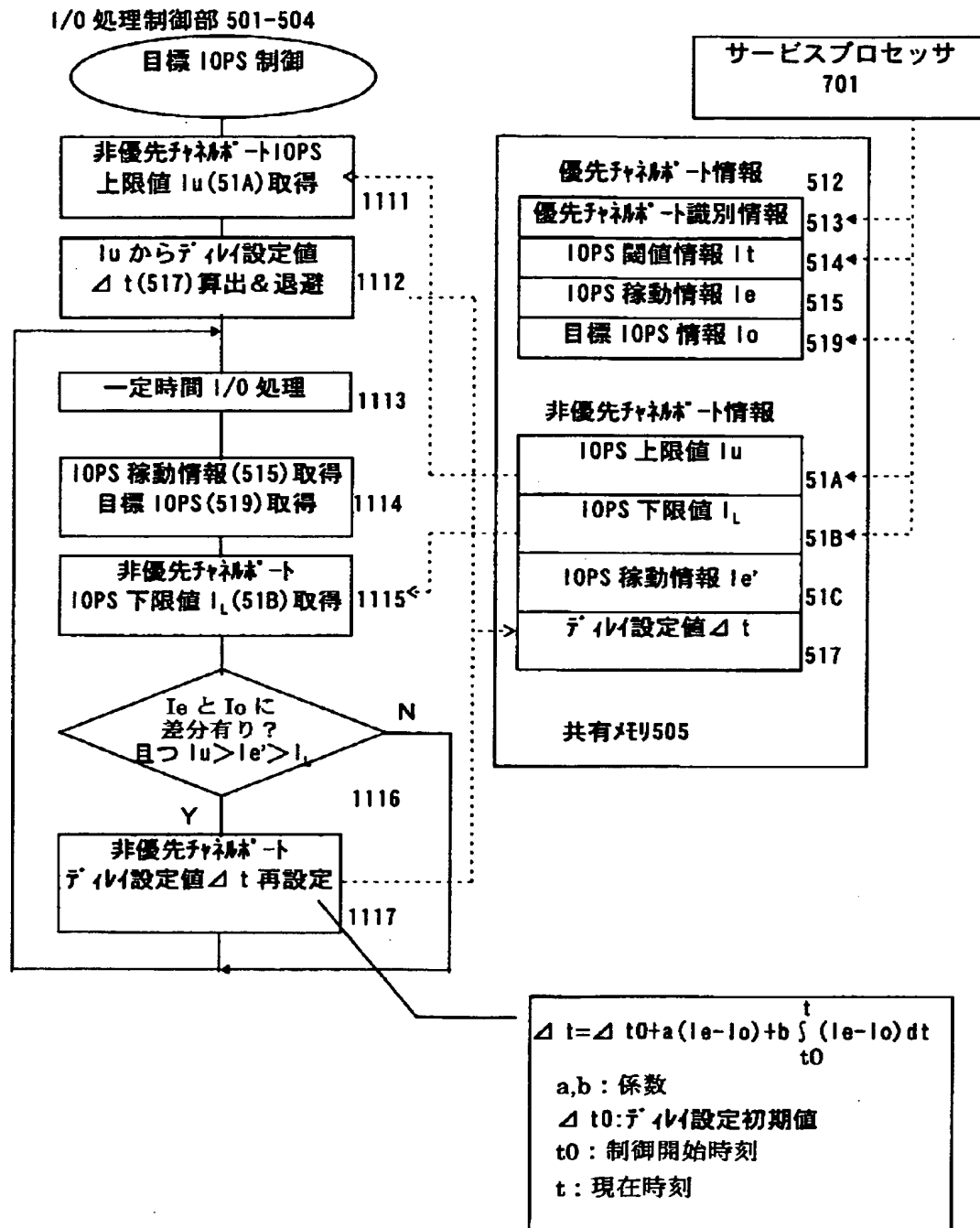
【図 1 0】

図 1 0



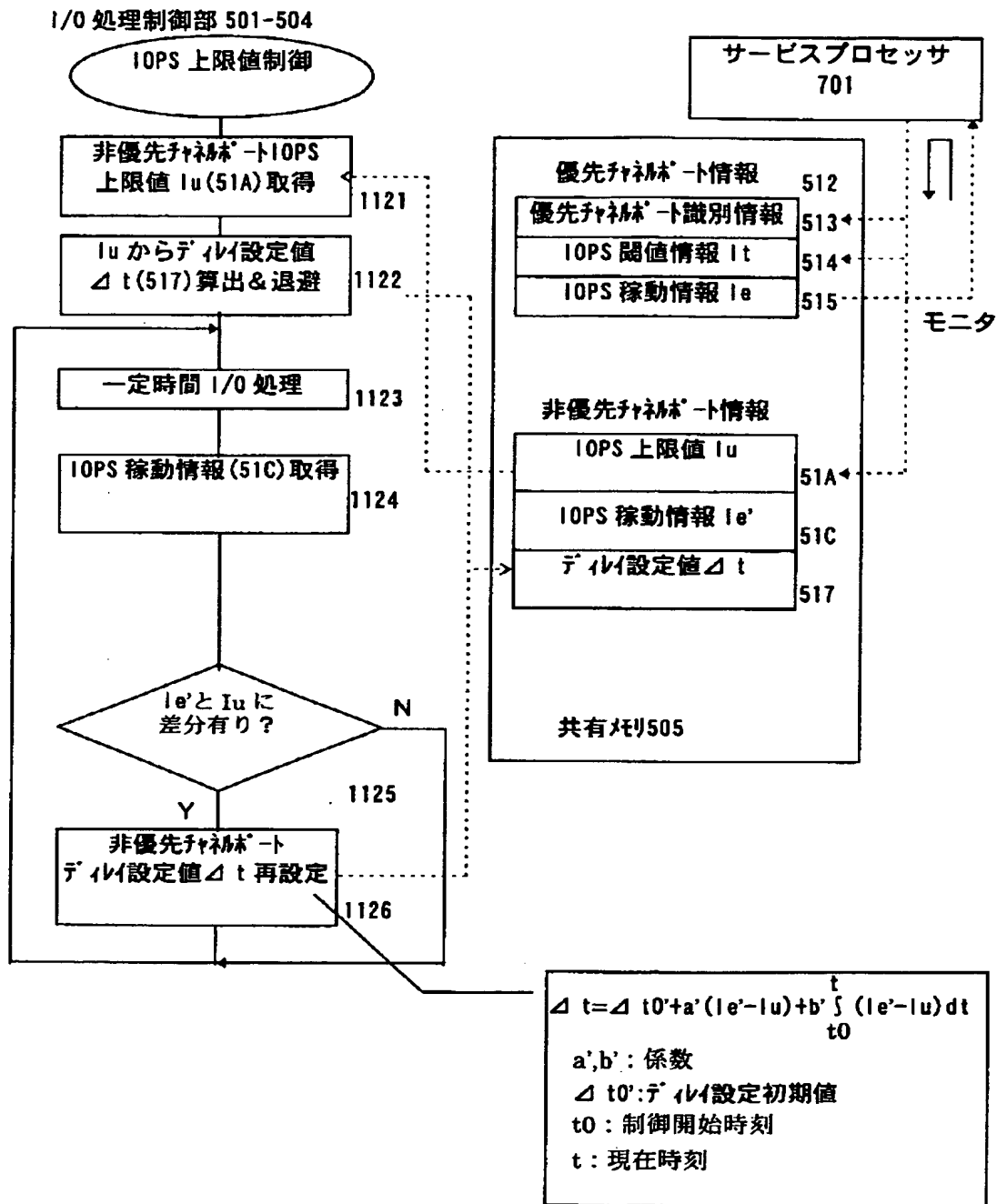
【図 11】

図 11



【図 1 2】

図 1 2



【書類名】 要約書

【要約】

【課題】

優先的に処理したい I / O を、他の I / O 処理に影響されることなく性能を維持しながら処理させるため、I / O 処理を優先と非優先に分け、非優先の I / O を抑制しながら処理することで、優先の I / O は他の I / O 処理に影響されることなく性能を維持しながら処理する。

【解決手段】

記憶制御装置 3 0 1 は I / O 処理制御部 5 0 1 ~ 5 0 4 に共通なメモリ 5 0 5 を保有し、メモリ 5 0 5 内に複数の I / O 処理を優先と非優先に分割し制御するための情報を管理し、非優先の I / O は処理を抑制しながら動作させる。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願 2000-300563
受付番号	50005047540
書類名	特許願
担当官	風戸 勝利 9083
作成日	平成12年10月 3日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000005108
【住所又は居所】	東京都千代田区神田駿河台四丁目6番地
【氏名又は名称】	株式会社日立製作所

【特許出願人】

【識別番号】	000233055
【住所又は居所】	神奈川県横浜市中区尾上町6丁目81番地
【氏名又は名称】	日立ソフトウェアエンジニアリング株式会社

【代理人】

【識別番号】	100075096
【住所又は居所】	東京都千代田区丸の内1-5-1 株式会社日立 製作所 知的所有権本部内
【氏名又は名称】	作田 康夫

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所

出 願 人 履 歴 情 報

識別番号 [000233055]

1. 変更年月日 1990年 8月 7日

[変更理由] 新規登録

住 所 神奈川県横浜市中区尾上町6丁目81番地

氏 名 日立ソフトウェアエンジニアリング株式会社